

(2) 000-293317 (P2000-293317A)

【特許請求の範囲】

【請求項1】 データ処理装置が直接アクセスを行う論理的記憶装置を実際にデータを記憶する物理的記憶装置に配置し、前記データ処理装置と前記物理的記憶装置の間のデータ転送を制御する記憶制御装置において、予め定めた指標に基づいて前記論理的記憶装置を前記物理的記憶装置に再配置すると共に再配置先の物理的記憶装置にデータを連続的に格納する論理的記憶装置再配置手段を有することを特徴とする記憶制御装置。

【請求項2】 データ処理装置が直接アクセスを行う論理的記憶装置と実際にデータを記憶する物理的記憶装置とを対応付け、前記データ処理装置と前記物理的記憶装置の間のデータ転送を制御する記憶制御装置において、前記データ転送の制御の運用中にデータ処理装置の論理的記憶装置へのアクセス情報を指標として採取するアクセス情報採取手段と、前記指標に基づいて前記論理的記憶装置を前記物理的記憶装置に再配置すると共に再配置先の物理的記憶装置にデータを連続的に格納する論理的記憶装置再配置手段とを有することを特徴とする記憶制御装置。

【請求項3】 請求項2に記載の記憶制御装置において、前記アクセス情報が、前記データ処理装置から前記論理的記憶装置へのアクセス頻度情報を含むことを特徴とする記憶制御装置。

【請求項4】 請求項2または請求項3に記載の記憶制御装置において、前記アクセス情報が、前記データ処理装置から前記論理的記憶装置へのアクセスパターン情報を含むことを特徴とする記憶制御装置。

【請求項5】 請求項1に記載の記憶制御装置において、前記指標が、前記論理的記憶装置に求められる信頼性であることを特徴とする記憶制御装置。

【請求項6】 請求項1から請求項5のいずれかに記載の記憶制御装置において、前記指標を保守員に提示する指標提示手段と、保守員からの再配置指示を受け付ける再配置指示受付手段とを具備したことを特徴とする記憶制御装置。

【請求項7】 請求項1から請求項5のいずれかに記載の記憶制御装置において、データ処理装置からの再配置指示を受け付ける再配置指示受付手段を具備したことを特徴とする記憶制御装置。

【請求項8】 請求項1から請求項5のいずれかに記載の記憶制御装置において、前記指標に基づいて再配置の要否を決定する再配置要否決定手段を具備したことを特徴とする記憶制御装置。

【請求項9】 請求項1から請求項8のいずれかに記載の記憶制御装置において、再配置中の論理的記憶装置にデータ処理装置からのアクセスがあったとき、再配置中の論理的記憶装置の再配置完了領域と再配置未完領域とを識別し、前記アクセス位置が前記再配置完了領域ならば再配置先の論理的記憶装置にアクセスさせ、前記アク

セス位置が前記再配置未完領域ならば当該論理的記憶装置にアクセスさせるアクセス位置切替手段をさらに具備したことを特徴とする記憶制御装置。

【請求項10】 データ処理装置がデータアクセスのために認識する論理的な記憶装置のデータを保持するための複数の物理的な記憶装置と、前記複数の物理装置とデータ処理装置との間のデータ転送を制御する記憶制御装置とを有する記憶装置システムにおいて、前記記憶制御装置は、前記データ処理装置によるアクセス状況を取得し、前記アクセス状況に基づいて前記論理的な記憶装置のデータを第1の物理的な記憶装置から第2の物理的な記憶装置に移動させることを特徴とする記憶装置システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、記憶制御装置に関し、さらに詳しくは、シーケンシャルアクセスの場合やランダムアクセスでヒット率が低い場合でもアクセス性能を向上することが出来る記憶制御装置およびデータの信頼性を向上することが出来る記憶制御装置に関する。特に、本発明は、ディスクアレイ向きの高機能ディスク装置、その高機能ディスク装置とディスク制御装置とにより構成される記憶装置サブシステム、およびその記憶装置サブシステムとデータ処理装置とにより構成される情報処理システムに有用である。

【0002】

【従来の技術】シカゴのイリノイ大学で開かれた「ACM SIGMOD」会議において発表された論文「D.Patterson,G. gibson, and R.H.Kartzt; A Case for Redundant Arrays of Inexpensive Disks (RAID), ACM SIGMOD Conference, Chicago, IL, (June 1988), pp.109-116」は、ディスクアレイ上のデータ配置に関する技術を開示している。

【0003】また、特開平7-84732号公報では、ディスク装置の一部をディスクキャッシュの如く用いる技術が開示されている。具体的には、ディスク装置を一時的にデータを格納するテンポラリ領域と最終的にデータを書き込む領域とに分け、更新データはパリティを生成せずに一旦テンポラリ領域に二重書きし、非同期にパリティ生成し、最終領域に書き込む。

【0004】一方、電気情報通信学会技術研究報告「DE95-68 (茂木他: Hot Mirroring を用いたディスクアレイのディスク故障時の性能評価、1995年12月、電気情報通信学会技報 Vol.95-No.407, pp.19-24)」には、アクセス頻度の違いにより、データを保持するRAIDレベルを動的に変更する技術が開示されている。具体的には、ディスク装置をRAID1構成の部分とRAID5構成の部分に分け、ライトアクセスのあったデータを優先的にRAID1構成の部分に格納するようにデータの格納位置を動的に変更することにより、アクセス頻度の高いデータはRAID1構成の部分に格

(3) 000-293317 (P2000-293317A)

納し、アクセス頻度の低いものはRAID5構成の部分に格納するように出来る。この技術によれば、記憶容量の異なる物理ディスク装置やRAIDレベルの異なる物理ディスク装置を記憶装置サブシステム内で混在させることが可能であり、論理ディスク装置内のデータを、そのアクセス頻度やアクセスパターンなどの指標に基づいて、任意の物理ディスク装置に格納することが出来る。また、アクセス頻度の高いデータを、より高速な物理ディスク装置に格納するように、動的に格納位置を変更することも出来る。なお、RAID1のディスクアレイは、データ処理装置からの書き込みデータに対して、その複製をミラーと呼ばれる副ディスク装置に書き込み、データの信頼性を確保する。冗長データが元のデータの複製であるため、冗長データ作成のオーバーヘッドが小さく、アクセス性能が良い。但し、物理的記憶装置の使用効率は、50%と低い。一方、RAID5のディスクアレイは、データ処理装置からの複数の書き込みデータに対して、パリティと呼ばれる冗長データを作成する。パリティ作成時に更新前データと更新前パリティのリードが必要なため、冗長データ作成のオーバーヘッドが大きく、アクセス性能は悪い。但し、複数のデータに対して1つのパリティを作成するため、記憶装置の使用効率はRAID1に比べ高い。

【0005】

【発明が解決しようとする課題】上記従来技術では、アクセスするデータ単位でデータの格納位置の変更を行うため、データ処理装置が直接アクセスを行う論理ディスク装置上では連続なデータが、実際にデータを記憶する物理ディスク装置上では非連続となってしまう。このため、一連のデータをリード/ライトするシーケンシャルアクセスの場合、実際には複数データをまとめてリード/ライトできなくなり、アクセス性能の低下を招く問題点がある。

【0006】一方、上記報告「DE95-68」の従来技術では、ライトの度に、アクセス頻度が低いと判断したデータをRAID1構成の部分からRAID5構成の部分に移し、空いたRAID1構成の部分にライトデータを書き込むため、アクセスパターンがランダムアクセスでヒット率が低い場合には、RAID1構成の部分に移したデータの多くは再びRAID5構成の部分に戻されることになる。このため、ヒット率が低い場合、アクセス性能の向上は期待できず、逆にデータを移す処理のオーバーヘッドがアクセス性能の低下を引き起こす問題点がある。

【0007】また、上記の従来技術では、データの信頼性の向上については全く考慮されていない問題点がある。

【0008】そこで、本発明の第1の目的は、シーケンシャルアクセスの場合やランダムアクセスでヒット率が低い場合でも、アクセス性能を向上することが出来る記

憶制御装置を提供することにある。また、本発明の第2の目的は、データの信頼性を向上することが出来る記憶制御装置を提供することにある。

【0009】

【課題を解決するための手段】第1の観点では、本発明は、データ処理装置が直接アクセスを行う論理的記憶装置を実際にデータを記憶する物理的記憶装置に配置し、前記データ処理装置と前記物理的記憶装置の間のデータ転送を制御する記憶制御装置において、予め定めた指標に基づいて前記論理的記憶装置を前記物理的記憶装置に再配置すると共に再配置先の物理的記憶装置にデータを連続的に格納する論理的記憶装置再配置手段を有することを特徴とする記憶制御装置を提供する。上記第1の観点による記憶制御装置では、アクセスするデータ単位でデータの格納位置の変更を行うのではなく、論理的記憶装置を単位として物理的記憶装置への再配置を行い、且つ、再配置先の物理的記憶装置にデータを連続的に格納する。従って、シーケンシャルアクセスの場合でも、アクセス性能を向上することが出来る。また、ライトの度にデータの格納位置の変更を行うのではなく、予め定めた指標に基づいて前記再配置を行うから、ランダムアクセスでヒット率が低い場合でも、アクセス性能を向上することが出来る。

【0010】第2の観点では、本発明は、データ処理装置が直接アクセスを行う論理的記憶装置と実際にデータを記憶する物理的記憶装置とを対応付け、前記データ処理装置と前記物理的記憶装置の間のデータ転送を制御する記憶制御装置において、前記データ転送の制御の運用中にデータ処理装置の論理的記憶装置へのアクセス情報を指標として採取するアクセス情報採取手段と、前記指標に基づいて前記論理的記憶装置を前記物理的記憶装置に再配置すると共に再配置先の物理的記憶装置にデータを連続的に格納する論理的記憶装置再配置手段とを有することを特徴とする記憶制御装置を提供する。上記第2の観点による記憶制御装置では、アクセスするデータ単位でデータの格納位置の変更を行うのではなく、論理的記憶装置を単位として物理的記憶装置への再配置を行い、且つ、再配置先の物理的記憶装置にデータを連続的に格納する。従って、シーケンシャルアクセスの場合でも、アクセス性能を向上することが出来る。また、ライトの度にデータの格納位置の変更を行うのではなく、アクセス情報を採取し、それを統計的に利用して前記再配置を行うから、ランダムアクセスでヒット率が低い場合でも、アクセス性能を向上することが出来る。

【0011】第3の観点では、本発明は、上記構成の記憶制御装置において、前記アクセス情報が、前記データ処理装置から前記論理的記憶装置へのアクセス頻度情報を含むことを特徴とする記憶制御装置を提供する。上記第3の観点による記憶制御装置では、アクセス頻度の高い論理的記憶装置をより高速な物理的記憶装置へ再配置

(4) 000-293317 (P2000-293317A)

することが出来る。従って、アクセス性能を向上することが出来る。

【0012】第4の観点では、本発明は、上記構成の記憶制御装置において、前記アクセス情報が、前記データ処理装置から前記論理的記憶装置へのアクセスパターン情報を含むことを特徴とする記憶制御装置を提供する。上記第4の観点による記憶制御装置では、シーケンシャルアクセスの比率の高い論理的記憶装置をよりシーケンシャルアクセス性能の高い物理的記憶装置へ再配置することが出来る。従って、アクセス性能を向上することが出来る。

【0013】第5の観点では、本発明は、上記構成の記憶制御装置において、前記指標が、前記論理的記憶装置に求められる信頼性であることを特徴とする記憶制御装置を提供する。上記第5の観点による記憶制御装置では、信頼性が高いことが求められる論理的記憶装置をより信頼性の高い物理的記憶装置へ再配置することが出来る。従って、データの信頼性を向上することが出来る。

【0014】第6の観点では、本発明は、上記構成の記憶制御装置において、前記指標を保守員に提示する指標提示手段と、保守員からの再配置指示を受け付ける再配置指示受付手段とを具備したことを特徴とする記憶制御装置を提供する。上記第6の観点による記憶制御装置では、保守員が再配置指示を入力できるため、非常に柔軟に前記再配置を行うことが出来る。

【0015】第7の観点では、本発明は、上記構成の記憶制御装置において、データ処理装置からの再配置指示を受け付ける再配置指示受付手段を具備したことを特徴とする記憶制御装置を提供する。上記第7の観点による記憶制御装置では、データ処理装置が再配置指示を入力できるため、保守員では判断不可能な高度の条件下で前記再配置を行うことが出来る。

【0016】第8の観点では、本発明は、上記構成の記憶制御装置において、前記指標に基づいて再配置の要否を決定する再配置要否決定手段を具備したことを特徴とする記憶制御装置を提供する。上記第8の観点による記憶制御装置では、記憶制御装置が再配置指示を自己決定するため、保守員やデータ処理装置に負担をかけなくて済む。

【0017】第9の観点では、本発明は、上記構成の記憶制御装置において、再配置中の論理的記憶装置にデータ処理装置からのアクセスがあったとき、再配置中の論理的記憶装置の再配置完了領域と再配置未完領域とを識別し、前記アクセス位置が前記再配置完了領域ならば再配置先の論理的記憶装置にアクセスさせ、前記アクセス位置が前記再配置未完領域ならば当該論理的記憶装置にアクセスさせるアクセス位置切替手段を具備したことを特徴とする記憶制御装置を提供する。上記第9の観点による記憶制御装置では、再配置中の論理的記憶装置の再配置完了領域と再配置未完領域とを識別し、データ処理

装置からのアクセス位置を切り替えるから、データ処理装置と物理的記憶装置の間のデータ転送を運用中に再配置を行うことが出来る。

【0018】第10の観点では、本発明は、データ処理装置がデータアクセスのために認識する論理的な記憶装置のデータを保持するための複数の物理的な記憶装置と、前記複数の物理装置とデータ処理装置との間のデータ転送を制御する記憶制御装置とを有する記憶装置システムにおいて、前記記憶制御装置は、前記データ処理装置によるアクセス状況を取得し、前記アクセス状況に基づいて前記論理的な記憶装置のデータを第1の物理的な記憶装置から第2の物理的な記憶装置に移動させることを特徴とする記憶装置システムを提供する。上記第10の観点による記憶制御装置では、データ処理装置によるアクセス状況に応じて論理的記憶装置のデータを第1から第2の物理的な記憶装置へと移動させるから、アクセス性能を向上することが出来る。

【0019】

【発明の実施の形態】以下、本発明の実施形態を説明する。なお、これにより本発明が限定されるものではない。

【0020】-第1の実施形態-

第1の実施形態は、各論理ディスク装置のアクセス情報を記憶制御装置で採取し、SVP（サービスプロセッサ）を通じて保守員に提示し、このアクセス情報に基づく保守員の再配置指示により、論理ディスク装置の物理ディスク装置への再配置を行うものである。

【0021】図1は、本発明の第1の実施形態にかかる記憶制御装置を含む情報処理システムのブロック図である。この情報処理システム1は、データ処理装置100と、記憶制御装置104と、1台以上の物理ディスク装置105と、SVP111とを接続してなっている。

【0022】前記データ処理装置100は、CPU101と、主記憶102と、チャンネル103とを有している。

【0023】前記記憶制御装置104は、1つ以上のディレクタ106と、キャッシュメモリ107と、ディレクトリ108と、不揮発性メモリ109と、不揮発性メモリ管理情報110と、論理物理対応情報300と、論理ディスク装置情報400と、アクセス情報500を有している。前記ディレクタ106は、データ処理装置100のチャンネル103と物理ディスク装置105の間のデータ転送、データ処理装置100のチャンネル103と前記キャッシュメモリ107の間のデータ転送および前記キャッシュメモリ107と物理ディスク装置105の間のデータ転送を行う。前記キャッシュメモリ107には、物理ディスク装置105の中のアクセス頻度の高いデータをロードしておく。このロード処理は、前記ディレクタ106が実行する。ロードするデータの具体例は、データ処理装置100のCPU101のアクセス対

(5) 000-293317 (P2000-293317A)

象データや、このアクセス対象データと物理ディスク装置105上の格納位置に近いデータ等である。前記ディレクトリ108は、前記キャッシュメモリ107の管理情報を格納する。前記不揮発性メモリ109は、前記キャッシュメモリ107と同様に、物理ディスク装置105の中のアクセス頻度の高いデータをロードしておく。前記不揮発性メモリ管理情報110は、前記不揮発性メモリ109の管理情報を格納する。前記論理物理対応情報300は、各論理ディスク装置(図2の200)が配置されている物理ディスク装置105上の位置および各物理ディスク装置105に配置されている論理ディスク装置(図2の200)を示す情報である。この情報を用いて、データ処理装置100のCPU101のアクセス対象データの物理ディスク装置105上の格納領域の算出などを行う。前記論理ディスク装置情報400は、各論理ディスク装置(図2の200)のアクセス可否等の状態を示す。前記アクセス情報500は、各論理ディスク装置(図2の200)のアクセス頻度やアクセスパターンなどの情報である。

【0024】論理物理対応情報300と論理ディスク情報400は、電源断などによる消失を防ぐために不揮発の媒体に記録する。

【0025】前記物理ディスク装置105は、データを記録する媒体と、記録されたデータを読み書きする装置とから構成される。

【0026】前記SVP111は、アクセス情報500の保守員への提示や、保守員からの再配置指示620の入力の受け付けを行う。また、保守員からの情報処理システム1への指示の発信や、情報処理システム1の障害状態等の保守員への提示を行う。

【0027】図2は、論理ディスク装置200と物理ディスク装置105の関連を表わした図である。論理ディスク装置200は、データ処理装置100のCPU101が直接アクセスする見掛け上のディスク装置で、アクセス対象データが実際に格納される物理ディスク装置105と対応している。論理ディスク装置200上のデータは、シーケンシャルアクセスを考慮して、物理ディスク装置105上に連続的に配置されている。論理ディスク装置200のデータが配置されている物理ディスク装置105がディスクアレイ構成の場合、該論理ディスク装置200は複数の物理ディスク装置105と対応する。また、物理ディスク装置105の容量が論理ディスク装置200より大きく、複数の論理ディスク装置のデータを1台の物理ディスク装置105に格納できる場合には、該物理ディスク装置105は複数の論理ディスク装置200と対応する。この論理ディスク装置200と物理ディスク装置105の対応は前記論理物理対応情報300で管理される。例えば、データ処理装置100のCPU101が論理ディスク装置200のデータ201をリードする時、記憶制御装置104で論理物理対応情

報300に基づき論理ディスク装置200に対応する物理ディスク装置105を求め、その物理ディスク装置105の領域内のデータ格納位置202を求め、データ転送を行う。

【0028】図3は、論理物理対応情報300を表わした図である。論理物理対応情報300は、論理ディスク構成情報310と、物理ディスク構成情報320とから構成される。前記論理ディスク構成情報310は、各論理ディスク装置200が配置されている物理ディスク装置105上の領域に関する情報であり、論理ディスク装置200から対応する物理ディスク装置105を求める時に用いる。一方、前記物理ディスク構成情報320は、各物理ディスク装置105に配置されている論理ディスク装置200に関する情報で、物理ディスク装置105から対応する論理ディスク装置200を求める時に用いる。

【0029】前記論理ディスク構成情報310は、物理ディスク装置グループ311、RAID構成312および開始位置313の組を、論理ディスク装置200の数だけ有している。前記物理ディスク装置グループ311は、当該論理ディスク装置200が配置されている物理ディスク装置105を示す情報である。前記RAID構成312は、前記物理ディスク装置グループ311のRAIDレベルを示す。前記開始位置313は、当該論理ディスク装置200が物理ディスク装置105上で配置されている先頭位置を示す。

【0030】前記物理ディスク構成情報320は、論理ディスク装置グループ321を、物理ディスク装置105の数だけ有している。前記論理ディスク装置グループ321は、当該物理ディスク装置105に配置されている論理ディスク装置200を示す。

【0031】図4は、論理ディスク情報400を表わした図である。論理ディスク情報400は、論理ディスク状態401と再配置完了ポインタ402とを、論理ディスク装置200の数だけ有している。前記論理ディスク状態401は、「正常」「閉塞」「フォーマット中」「再配置中」などの論理ディスク装置200の状態を表わす。前記再配置完了ポインタ402は、前記論理ディスク状態401が「再配置中」の時のみ有効な情報で、当該論理ディスク装置200の再配置処理を完了している領域の次の位置すなわち当該論理ディスク装置200が未だ再配置処理を終えていない領域の先頭位置を示す。「再配置中」におけるデータアクセス時、再配置完了ポインタ402よりも前の領域へのアクセスの場合には、再配置後の物理ディスク装置105へアクセスしなければならない。一方、再配置完了ポインタ402以後の領域へのアクセスの場合には、再配置前の物理ディスク装置105へアクセスしなければならない。

【0032】図5は、アクセス情報500を表わしている。アクセス情報500は、アクセス頻度情報501と

(6) 000-293317 (P2000-293317A)

アクセスパターン情報502とを、論理ディスク装置200の数だけ有している。このアクセス情報500は、記憶制御装置104、データ処理装置100、SVP111のいずれからも参照することが出来る。前記アクセス頻度情報501は、単位時間あたりの当該論理ディスク装置200へのアクセス回数を管理する。このアクセス頻度情報501は、各論理ディスク装置200の中でアクセス頻度の高いもの又は低いものを求める指標として用いる。前記アクセスパターン情報502は、当該論理ディスク装置200へのシーケンシャルアクセスとランダムアクセスの割合を管理する。このアクセスパターン情報502は、シーケンシャルアクセスが多く、よりシーケンシャル性能の高い物理ディスク装置105に再配置するのが望ましい論理ディスク装置200を求める指標として用いる。

【0033】次に、記憶制御装置104の動作を説明する。図6は、記憶制御装置104の動作を詳細に表わした図である。まず、リード/ライト処理時の動作について説明する。ディレクタ106は、通常リード/ライト処理を実行する際、CPU101からチャンネル103を経由してCPUからの指示600を受け取る。このCPUからの指示600は、リード（またはライト）対象のレコードが記憶されている論理ディスク装置200を指定する指定情報1と、リード（またはライト）対象のレコードが記憶されている論理ディスク装置200内の位置（トラック、セクタ、レコード）を指定する指定情報2とを含んでいる。ディレクタ106は、物理ディスク装置上のアクセス位置算出処理（610）で、前記CPUからの指示600と論理物理対応情報300とを用いて、物理ディスク装置105上でのアクセス位置を算出する。この物理ディスク装置アクセス位置算出処理（610）については図8を参照して後で詳述する。その後、たとえばリード処理では、算出した物理ディスク装置105上のデータ格納位置202のデータをキャッシュメモリ107上に読み上げてデータ201とし、その読み上げたデータ201をチャンネル103を通じて主記憶102に転送する。

【0034】次に、アクセス情報500の採取処理について説明する。CPU101からのリード/ライト処理のアクセス時に、ディレクタ106は、アクセス対象論理ディスク装置200のアクセス情報500を更新する。アクセス頻度情報501の採取は、例えば、アクセスの度に内部カウンタをカウントアップしていき、一定時間または一定回数のアクセス経過後のアクセス時に、前記内部カウンタからアクセス頻度を判定する。アクセスパターン情報502の採取は、例えば、アクセスの度に内部カウンタにシーケンシャルアクセス回数をカウントアップしていき、一定時間または一定回数のアクセス経過後のアクセス時に、前記内部カウンタからアクセスパターンを判定する。

【0035】次に、再配置指示620を説明する。保守員は、SVP111を通じて提示されたアクセス情報500を参照して、各論理ディスク装置200の再配置の必要性を検討する。この検討の結果、再配置を決定した論理ディスク装置200があれば、SVP111を通じて記憶制御装置104に対して再配置指示620を出す。この再配置指示620は、再配置対象の論理ディスク装置200を2つ指定する指示情報1-2からなる。保守員が行う検討の内容は、後述する第3の実施形態で図10を参照して説明する論理ディスク装置再配置要否決定処理（910）と同様である。

【0036】次に、論理ディスク装置再配置処理（630）を説明する。ディレクタ106は、前記再配置指示620を受けて、指定された2つの論理ディスク装置200の間で論理ディスク装置再配置処理（630）を行う。図7は、論理ディスク装置再配置処理部630の処理フロー図である。ステップ700では、論理ディスク情報400のうちの指定された2つの論理ディスク装置200の論理ディスク状態401を「再配置中」に設定する。ステップ701では、論理ディスク情報400のうちの指定された2つの論理ディスク装置200の再配置完了ポインタ402を各論理ディスク装置200の先頭位置に初期化する。ステップ702では、論理ディスク情報400のうちの指定された2つの論理ディスク装置200の再配置完了ポインタ402をチェックし、全領域の再配置が完了していなければステップ703へ進み、完了していればステップ707へ進む。

【0037】ステップ703では、再配置完了ポインタ402が示すデータ位置から再配置処理の1回の処理単位分のデータに対して物理ディスク装置105からキャッシュメモリ107上へのデータ転送を行う。ここで、1回の処理単位分のデータ量は、再配置対象の2つの論理ディスク装置200の冗長データ1つに対応する各データ量の最小公倍数に決定される。たとえば、再配置をRAID5の論理ディスク装置200とRAID1の論理ディスク装置200の間で行うならば、RAID1の論理ディスク装置200の冗長データ1つに対応するデータ量は“1”であるから、1回の処理単位分のデータ量は、RAID5の論理ディスク装置200の冗長データ1つに対応するデータ量すなわちパリティ1つに対応するデータ量に決定される。

【0038】ステップ704では、再配置対象の各論理ディスク装置200の再配置先論理ディスク装置200がパリティを有するRAIDレベルのものである場合、キャッシュメモリ107上の再配置対象の1回の処理単位分のデータ201に対してパリティを生成する。ステップ705では、キャッシュメモリ107上の再配置対象の1回の処理単位分のデータ201および前記ステップ704で作成したパリティを、再配置先の物理ディスク装置105へ書き込む。ステップ706では、1回の

!(7) 000-293317 (P2000-293317A)

処理単位分だけ再配置完了ポイント402を進める。そして、前記ステップ702に戻る。

【0039】なお、上記ステップ703、704において、データおよびパリティは、不揮発性メモリ109にも転送して二重化し、キャッシュ障害によるデータ消失を防ぐ。この理由は、上記ステップ705での書き込み時に、例えば、第1の論理ディスク装置200と第2の論理ディスク装置200のデータのうち、第1の論理ディスク装置200のデータを物理ディスク装置105

(元は第2の論理ディスク装置200に配置されていた物理ディスク装置105)へ書き込んだ段階で障害によりキャッシュメモリ107上のデータがアクセス不能になったとすると、書き込みが終了してない第2の論理ディスク装置200のデータが消失するからである(元は第2の論理ディスク装置200に配置されていた物理ディスク装置105には、上記のように第1の論理ディスク装置200のデータが上書きされてしまっている)。

【0040】ステップ707では、論理物理対応情報300を更新する。すなわち、論理ディスク構成情報310と物理ディスク構成情報321を変更する。ステップ708では、論理ディスク情報400の論理ディスク状態401を元の状態に戻し、再配置処理(630)を終了する。

【0041】次に、物理ディスク装置アクセス位置算出処理(610)を説明する。図8は、物理ディスク装置アクセス位置算出処理部610の処理フロー図である。ステップ800では、論理ディスク情報400のうちのアクセス対象論理ディスク装置200の論理ディスク状態401が「再配置中」であるか否かをチェックし、「再配置中」ならばステップ801に進み、「再配置中で」なければステップ803に進む。

【0042】ステップ801では、論理ディスク情報400のうちのアクセス対象論理ディスク装置200の再配置完了ポイント402とアクセスデータ位置とを比較し、アクセスデータ位置が再配置完了ポイント402の指す位置以後ならばステップ802に進み、アクセスデータ位置が再配置完了ポイント402の指す位置より前ならばステップ803に進む。

【0043】ステップ802では、当該論理ディスク装置200の再配置先の論理ディスク装置200をアクセス対象にする。そして、ステップ804へ進む。

【0044】ステップ803では、当該論理ディスク装置200をアクセス対象とする。

【0045】ステップ804では、アクセス対象の論理ディスク装置200に対応した物理ディスク装置105上でのアクセス位置を、論理物理対応情報300を用いて算出する。

【0046】以上の第1の実施形態にかかる情報処理システム1および記憶制御装置104によれば、アクセス情報500に基づく保守員の判断により、アクセス頻度

の高い論理ディスク装置をより高速な物理ディスク装置へ再配置することが出来る。また、シーケンシャルアクセスの比率の高い論理ディスク装置をよりシーケンシャルアクセス性能の高い物理ディスク装置へ再配置することが出来る。従って、アクセス性能を向上することが出来る。

【0047】-第2の実施形態-

上記第1の実施形態を変形して、記憶制御装置104からアクセス情報500をデータ処理装置100に提示し、データ処理装置100が再配置要否を決定し記憶制御装置104に再配置指示(620相当)を出すようにしてもよい。

【0048】-第3の実施形態-

第3の実施形態は、再配置指示をSVP111やデータ処理装置100から受けるのではなく、記憶制御装置104が自己決定するものである。

【0049】図9は、記憶制御装置104の動作を詳細に表わした図である。第1の実施形態(図6)との違いは、論理ディスク再配置要否決定処理部910が再配置指示620を出すことである。

【0050】図10は、上記論理ディスク再配置要否決定処理部910の処理フロー図である。この論理ディスク再配置要否決定処理(910)は、ディレクタ106が一定周期で各論理ディスク装置200のアクセス情報500を検査して行う。ステップ1000では、アクセス情報500のアクセス頻度情報501を参照し、アクセス頻度が規定値を超え且つ配置されている物理ディスク装置105が比較的低速なものである論理ディスク装置(以下、これを第1候補論理ディスク装置という)200があるか否かをチェックし、該当する論理ディスク装置200があればステップ1001へ進み、なければステップ1005へ進む。

【0051】ステップ1001では、前記第1候補論理ディスク装置200のアクセスパターン情報502を参照し、シーケンシャルアクセスの比率が規定値以上であるか否かをチェックし、規定値以上でなければステップ1002へ進み、規定値以上であればステップ1004へ進む。

【0052】ステップ1002では、前記第1候補論理ディスク装置200より高速な物理ディスク装置105に配置されている論理ディスク装置200のアクセス頻度情報501を参照し、アクセス頻度が規定値以下の論理ディスク装置(以下、これを第2候補論理ディスク装置という)200があるか否かをチェックし、あればステップ1003へ進み、なければステップ1005へ進む。

【0053】ステップ1003では、前記第1候補論理ディスク装置200と前記第2候補論理ディスク装置200の間で再配置処理(630)が必要であると決定し、再配置指示620を出す。そして、処理を終了す

!(8) 000-293317 (P2000-293317A)

る。

【0054】ステップ1004では、前記第1候補論理ディスク装置200よりシーケンシャル性能の高い物理ディスク装置105に配置されている論理ディスク装置200のアクセスパターン情報502を参照し、シーケンシャルアクセスの比率が規定値以下の論理ディスク装置（以下、これを第2候補論理ディスク装置という）200があるか否かをチェックし、あれば前記ステップ1003へ進み、なければ前記ステップ1002へ進む。

【0055】ステップ1005では、論理ディスク装置200の再配置処理（630）は不要であると決定する。そして、処理を終了する。

【0056】以上の第3の実施形態にかかる情報処理システム1および記憶制御装置104によれば、アクセス情報500に基づいて自動的に、アクセス頻度の高い論理ディスク装置をより高速な物理ディスク装置へ再配置することが出来る。また、シーケンシャルアクセスの比率の高い論理ディスク装置をよりシーケンシャルアクセス性能の高い物理ディスク装置へ再配置することが出来る。従って、アクセス性能を向上することが出来る。

【0057】－第4の実施形態－

上記第1～第3の実施形態を変形して、アクセス情報500に代えて又は加えて、論理ディスク装置200に要求される信頼性を再配置処理要否決定の指標に用いてもよい。信頼性を指標に用いれば、論理ディスク装置200上のデータの信頼性を向上させることが出来る。

【0058】

【発明の効果】本発明の記憶制御装置によれば、シーケンシャルアクセスの場合やランダムアクセスでヒット率が低い場合でも、アクセス性能を向上することが出来る。また、本発明の記憶制御装置によれば、データの信頼性を向上することが出来る。

【図面の簡単な説明】

【図1】本発明の第1の実施形態にかかる記憶制御装置を含む情報処理システムのブロック図である。

【図2】論理ディスク装置と物理ディスク装置との対応関係の説明図である。

【図3】論理物理対応情報の構成例示図である。

【図4】論理ディスク情報の構成例示図である。

【図5】アクセス情報の構成例示図である。

【図6】本発明の第1の実施形態における記憶制御装置の動作を示すブロック図である。

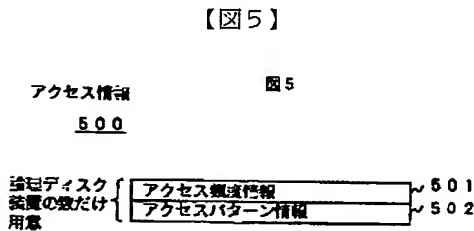
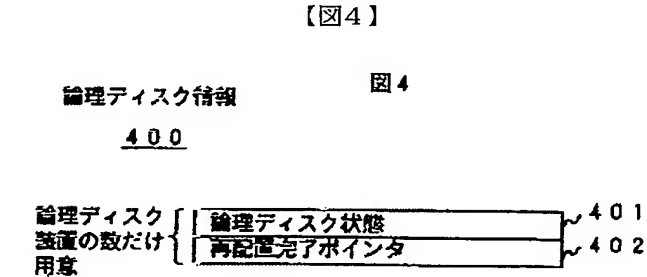
【図7】論理ディスク装置再配置処理部の処理フロー図である。

【図8】物理ディスク装置アクセス位置算出処理部の処理フロー図である。

【図9】本発明の第3の実施形態における記憶制御装置の動作を示すブロック図である。

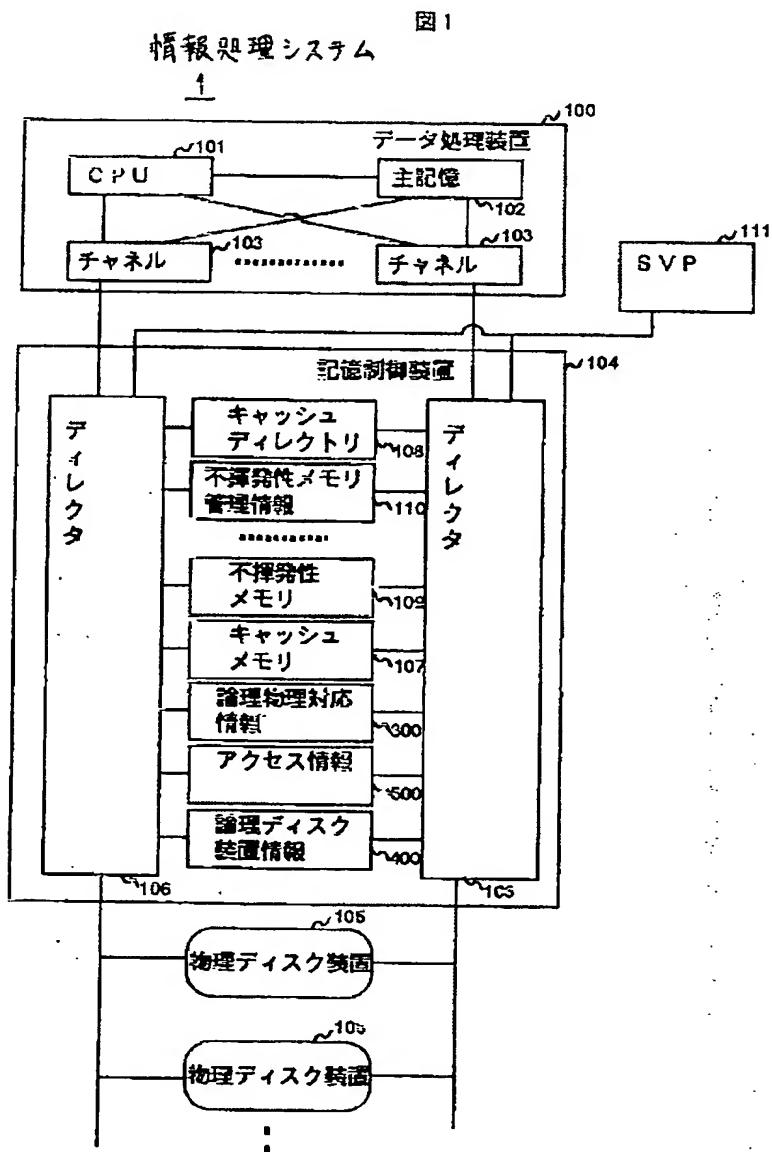
【図10】論理ディスク装置再配置要否決定処理部の処理フロー図である。

- 【符号の説明】
- 1 …情報処理システム
 - 100…データ処理装置
 - 101…CPU
 - 102…主記憶
 - 103…チャンネル
 - 104…記憶制御装置
 - 105…物理ディスク装置
 - 106…ディレクタ
 - 107…キャッシュメモリ
 - 108…キャッシュディレクトリ
 - 109…不揮発性メモリ
 - 110…不揮発性メモリ管理情報
 - 111…SVP
 - 200…論理ディスク装置
 - 201…データ
 - 202…データ格納位置
 - 300…論理物理対応情報
 - 400…論理ディスク情報
 - 500…アクセス情報
 - 600…CPUからの指示
 - 610…物理ディスク装置上のアクセス位置算出処理部
 - 620…指示情報
 - 630…論理ディスク装置再配置処理部
 - 910…論理ディスク再配置要否決定処理部



(9) 000-293317 (P2000-293317A)

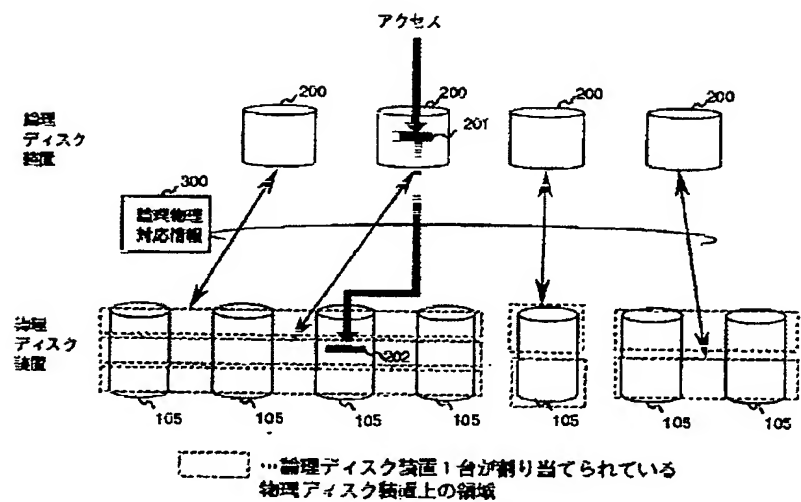
【図1】



(特 0) 100-293317 (P2000-293317A)

【図2】

図 2



【図3】

図 3

論理物理対応情報
300

・論理ディスク構成情報 310

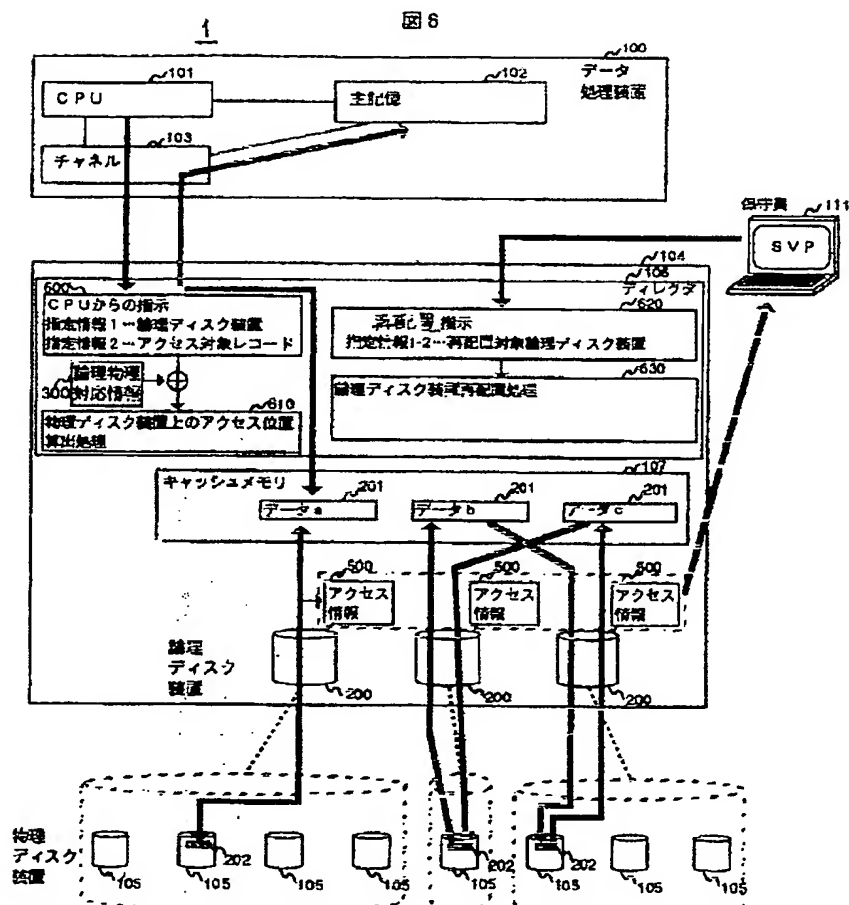
論理ディスク 装置の数だけ 用意	物理ディスク装置グループ	311
	RAID構成	312
	開始位置	313

・物理ディスク構成情報 320

物理ディスク 装置の数だけ 用意	論理ディスク装置グループ	321
------------------------	--------------	-----

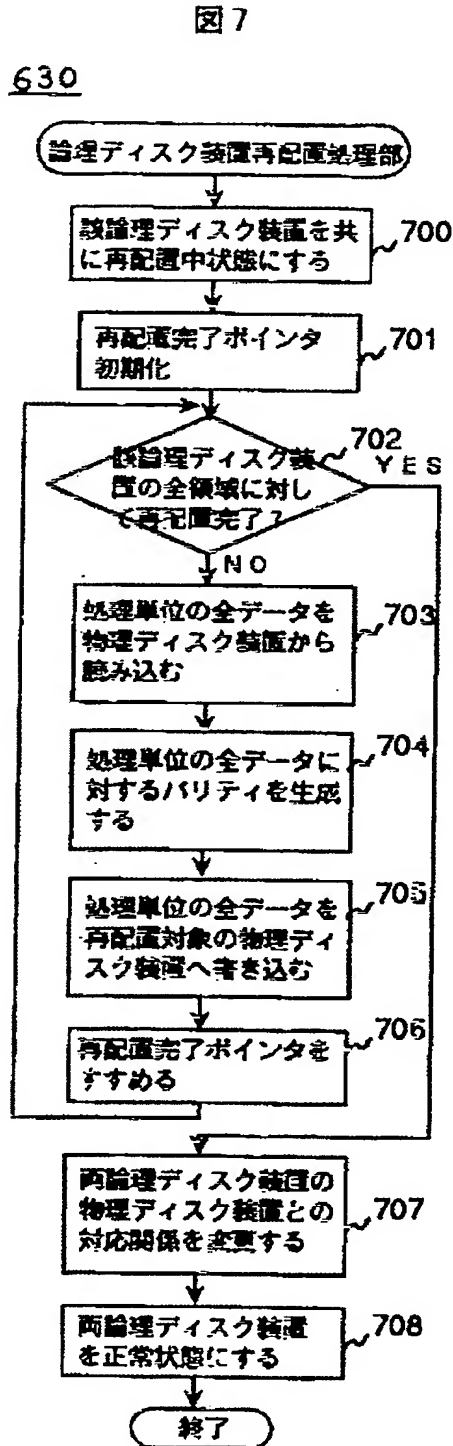
(図1) 100-293317 (P2000-293317A)

【図6】

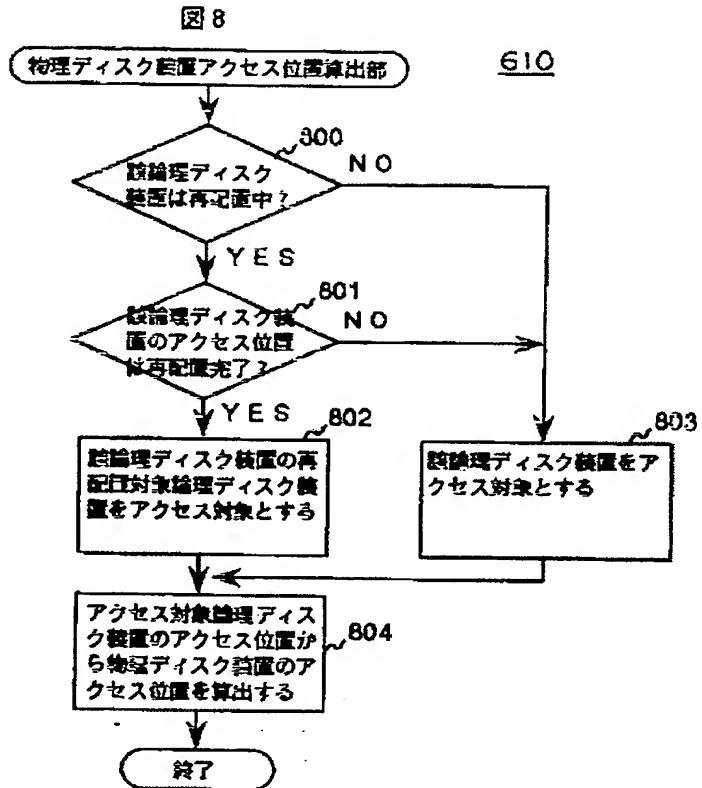


(2) 100-293317 (P2000-293317A)

【図7】

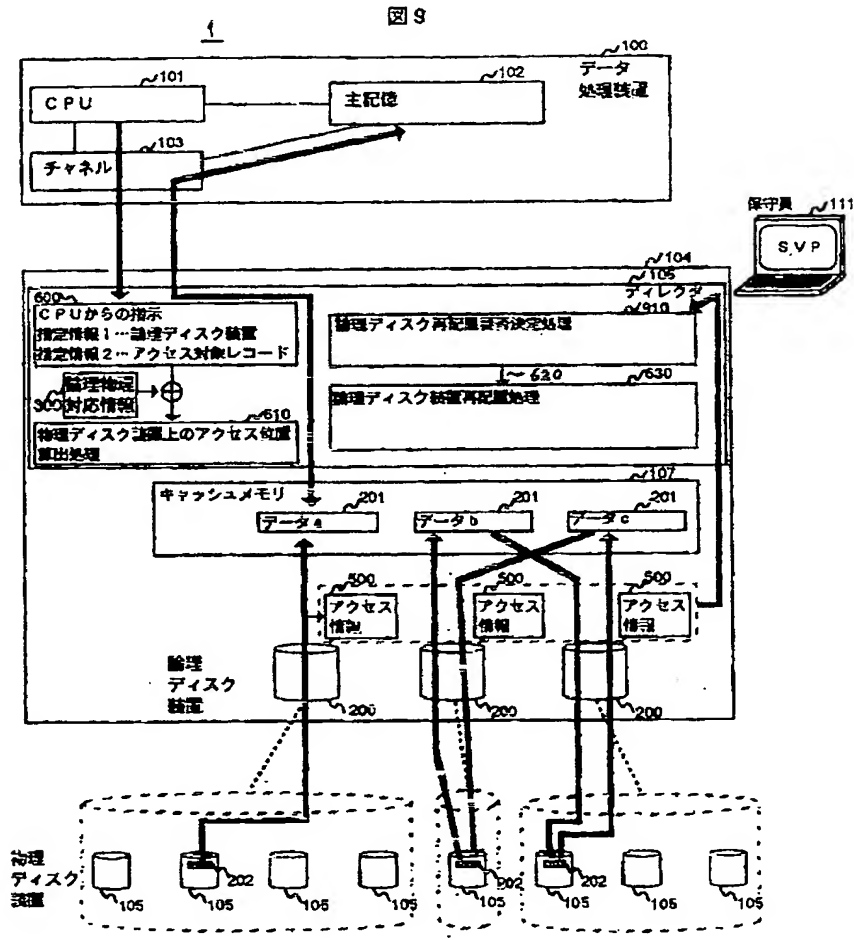


【図8】



(第3) 100-293317 (P2000-293317A)

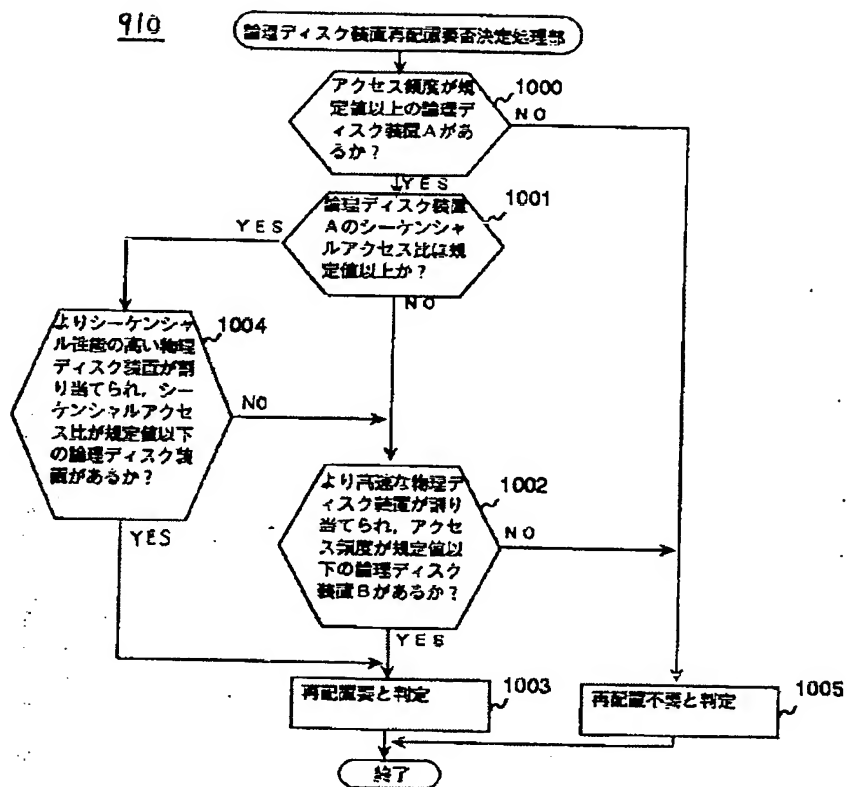
【図9】



(株) 00-293317 (P2000-293317A)

【図10】

図10



フロントページの続き

(51)Int.Cl.⁷
G11B 20/12

識別記号

FI
G11B 20/12

(参考)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)